

## SOUND SYNTHESIZER

Patent number: JP56051795  
Publication date: 1981-05-09  
Inventor: SAGAYAMA SHIGEKI; ITAKURA FUMITADA  
Applicant: NIPPON TELEGRAPH & TELEPHONE  
Classification:  
- international: G10L1/00  
- european:  
Application number: JP19790128364 19791003  
Priority number(s): JP19790128364 19791003

Abstract not available for JP56051795

Data supplied from the esp@cenet database – Worldwide

⑨ 日本国特許庁 (JP)

⑩ 特許出願公開

⑫ 公開特許公報 (A)

昭56—51795

⑤ Int. Cl.<sup>3</sup>  
G 10 L 1/00

識別記号

庁内整理番号  
7350—5D

⑬ 公開 昭和56年(1981)5月9日

発明の数 1  
審査請求 有

(全 6 頁)

⑭ 音声合成装置

⑮ 特 願 昭54—128364

⑯ 出 願 昭54(1979)10月3日

⑰ 発 明 者 嵯峨山茂樹

武蔵野市緑町3丁目9番11号日

本電信電話公社武蔵野電気通信

研究所内

⑱ 発 明 者 板倉文忠

武蔵野市緑町3丁目9番11号日

本電信電話公社武蔵野電気通信

研究所内

⑲ 出 願 人 日本電信電話公社

⑳ 代 理 人 弁理士 草野卓

明 細 書

1 発明の名称 音声合成装置

2 特許請求の範囲

(1) 合成しようとする音声に応じた各正弦波成分の周波数、振幅及びビッチ周期を数ミリ秒ないし数十ミリ秒ごとに指定する手段と、上記指定されたそれぞれの周波数及び振幅を持つ複数の正弦波を発生する正弦波発生手段と、その正弦波の発生を上記指定されたビッチ周期ごとに位相を初期化する初期化手段と、それ等初期化された正弦波の和を合成して音声信号を得る合成手段とを具備する音声合成装置。

3 発明の詳述な説明

この発明は合成音声を生ずる装置に関するものである。

音声合成装置は各種の通信装置、例えば音声応答装置、音声符号化復号化装置、音声出力つきの通信端末装置などや民生機器例えば教育機器、各種家庭電気器具などやその他の人間の声を発生する各種の応用分野における音声合成の利用が考え

られる。

従来の音声合成の方式として最も有力なものは PARCOR 形音声合成方式 (特許第 876024 号「音声合成装置」) あるいは LSP 形音声合成方式 (特許第 54-022318「音声合成器」) などである。特に前者の方式は最近 LSI 化され広く応用されるきざしが見えている。しかしこれらの従来の音声合成の方式には次のような欠点があつた。

例えば PARCOR 形音声合成方式は音声のスペクトル形状を PARCOR 係数と呼ばれる 8 ～ 12 個程度のパラメータにより表現するのであるが、これらのパラメータは音声現象との対応がつけにくく誤りが面倒であり、例えば規則による合成 (ルール合成) の場合この性質が障害となる。また量子化に関する性質が複雑で音声情報の圧縮率について問題を持つ。このために前記 LSP 音声合成方式が発明されたのであるが、この方式はいまだ装置の価格が高い。

いずれにせよこれらの方式は閉ループを持つ

(1)

(2)

繰形フィルタ(リカージブフィルタ recursive filter)を用いるため、発振や不安定性などの問題を内在している。このため例えば上記の音声合成方式をアナログ演算回路により実現することは要求される安定性や精度などの点からかなり難しく、一般にデジタル回路によつて実現する方が有利と考えられている。しかしデジタル回路を用いて実現する場合も演算経路を比較的長くする必要があつたり、パラメータの量子化に細心の注意を必要としたり種々の問題があつた。またそのため装置のハードウェア化も一般に複雑となり価格が高くなつていた。

この発明はこれらの欠点を根本的になくし音声合成フィルタを必要とせず、従つて比較的簡単なハードウェアで実現でき、またパラメータの量子化に関しても比較的短く微差演算経路を短くすることができ、また合成音の品質も十分実用に耐え得、広い応用分野へ適用することができる音声合成装置を提供するものである。

まずこの発明による音声合成の原理を説明する。

(3)

通れ  $i = 0, 1, \dots, 2n-1$  について等しくなるように  $\omega_1, \dots, \omega_n, A_1, \dots, A_n$  を決めることができ、これらを音声のフレームごとの分析結果とするものである。これは一種の音声スペクトル符号化と考えてよい。上述の文献では解の存在 ( $0 \leq \omega_i \leq \pi, A_i > 0, i = 1, \dots, n$ ) の条件や具体的な能率的な解法についても詳しく述べている。

この発明の音声合成の原理はパラメータ  $\{\omega_1, \dots, \omega_n, A_1, \dots, A_n\}$  を与えて(1)式により合成音声  $\{x_i\}$  を各時刻  $i$  についてつぎつぎと計算するものである。しかしこれだけでは音声としては聞えない。なぜならば音声は有声音と無声音に分類されるが、有声音(母音や濁子音など)は周期的な信号であるが(この周期をビツタと言う)、上記のモデルにはその周期性がない、つまりビツタの値がないからである。そのためビツタ周期ごと(1)式における時刻  $i$  をゼロにリセットし、つまり位相初期化する。これがこの発明のキーポイントの一つである。無声音についてはビツタは存

(5)

在しないが(1)式のままではスペクトルが日本の横スペクトルとなつて音声信号の真価とは異なるためランダムなビツタ周期を与えてランダムに位相初期化することによりスペクトルを広げる。

$$x_i = A_1 \sin \omega_1 i + \dots + A_n \sin \omega_n i \quad (1)$$

のようにモデル化する。但し  $i$  は離散的な時刻を扱わず連続、 $n$  は正弦波成分の数(通常4~6個程度)、 $\omega_i$  は第  $i$  正弦波成分の角周波数( $0 \leq \omega_i \leq \pi$ )、 $A_i$  は第  $i$  正弦波成分の振幅である。このモデルを複合正弦波モデルと呼ぶ。このモデルによる音声分析法については日本音声学会音声研究会資料87.9-06(1979年5月)豊城山・敬重「複合正弦波モデルによる音声の分析」で詳しく述べられてある。この分析法は(1)式(モデル)の自己相関関数と実際の音声の自己相関関数とが

(4)

以上を要する。電圧制御正弦波発生器  $1_1 \sim 1_n$  が設けられそれぞれ端子  $2_1 \sim 2_n$  を通じて与えられる角周波数  $\omega_1 \sim \omega_n$  に対応する電圧が電圧制御正弦波発生器  $1_1 \sim 1_n$  に入力され、振幅  $A_1$ 、角周波数  $\omega_1 \sim \omega_n$  の正弦波をそれぞれ出力する。これら発生器  $1_1 \sim 1_n$  はリセット端子を持ちこれにパルスを与えると正弦波出力の位相はゼロから再スタートする。ビツタ周期(あるいは周波数)に対応する電圧が端子  $1_1$  を通じて電圧制御パルス発生器  $1_2$  に与えられ、このパルス発生器  $1_2$  からそのビツタ周期ごとにパルスが発生される。このビツタ周期パルスにより電圧制御正弦波発生器  $1_1 \sim 1_n$  はリセットされる。電圧制御正弦波発生器  $1_1 \sim 1_n$  の出力は電圧制御可変利得増幅器  $3_1 \sim 3_n$  に供給されて振幅  $A$  の正弦波に端子  $4_1 \sim 4_n$  より

(6)

振幅  $A_1 \sim A_n$  がそれぞれ乗算される。これ等乗算出力は加算器 13 で合成されて出力端子 14 へ供給される。

動作は上述の原理に基づき合成しようとする音と対応して数  $m$  のないし数十  $m$  のごとく端子  $2_1 \sim 2_n, 4_1 \sim 4_n$  及び 11 に与える角周波数、振幅、ビツテ周波にそれぞれ対応した入力電圧を変化させてゆくとそれに対応した合成音声が出力から得られる。

第1図は各正弦波成分の振幅を直接与えてそれらの和を出力としたものであるが、一般に音声のダイナミツクレンジは広いため、各正弦波成分の振幅 ( $A_1 \dots A_n$ ) も大きなダイナミツクレンジを持つ。実効的検測によれば(1)式を

$$x_i = A \{ a_1 \sin \omega_1 t + \dots + a_n \sin \omega_n t \} \quad (2)$$

但し  $a_1^2 + \dots + a_n^2 = 1, 0 < a_i < 1 (i=1 \dots n)$  のように書き改めし ( $A_1 \dots A_n$ ) のかわりに ( $A, a_1 \dots a_n$ ) を与えた方が音声合成パラメータの情報圧縮符号化に有利である。今波 ( $a_1 \dots a_n$ ) を各正弦波成分の係数、 $A$  を利得と呼ぶこ

(7)

とにする。各正弦波成分の周波数と係数、利得及びビツテ周波を与えて音声合成をする場合のこの典型的例を第2図に示す。ここでは  $n=5$  の場合について示している。第2図中第1図と対応する部分には同一符号を付けてある。この例では端子  $4_1 \sim 4_5$  からそれぞれ係数  $a_1 \sim a_5$  を示す信号を可変利得増幅器  $3_1 \sim 3_5$  へ供給する。加算器 13 の出力はレベル制御器 15 で端子 16 よりの利得  $A$  を示す信号と乗算されて出力される。可変利得増幅器  $3_1 \sim 3_5$  の代わりに乗算器あるいは可変減衰率減衰器を用いてもよい。レベル制御器 15 としては乗算器あるいは可変利得増幅器または可変減衰率減衰器を用いることができる。

以上述べた装置においてはビツテ周波ごとく位相を初期化しているが、本質的には正弦波の和のみで合成音声を作り出そうとしていた。しかし実際の音声信号は数本の線スペクトルで表わせるのではなく、一般に連続スペクトルである。そこで線スペクトルの幅を広げるためにビツテ周波に同期した包絡線パターンを乗じることが好ましい。

(8)

第3図にその例を示し電圧制御パルス発生器 12 のパルス出力が包絡線発生器 17 のトリガ入力に与えられ、このパルスに同期して振幅が端子 16 の利得電圧に比例した包絡線信号が発生される。包絡線信号としては具体的には指数関数やレイズドコサイン関数 ( $(\exp t + 1)/2$ ) が用いられる。この包絡線信号が乗算器 15 へ供給される。

以上に述べた装置においては角周波数 ( $\omega_i$ )、係数 ( $a_i$ )、利得  $A$  及びビツテ周波などの制御入力は数  $m$  のないし数十  $m$  のごとくに変化させてゆくことを前提にしていたが、これらの入力は階段状に変化させるより滑らかに変化させた方が良好な合成音声が得られることが実験的にわかっている。そのため各正弦波成分の周波数の制御電圧及び振幅(あるいは係数と利得)の制御電圧をなめらかに変化させるためのフィルタ(一種の低域通過フィルタ)に通してから与えるのがよい。第4図はその例を示し平滑化フィルタ 21, 22 ( $i=1, 2 \dots n$ ) により階段状に変化する端子  $2_i, 4_i$  の角周波数  $\omega_i$ 、係数  $a_i$  の制御電圧をそれぞれ

(9)

滑らかにする。

次に以上説明した装置の各部の詳細例を述べる。第5図は発振周波数が電圧により制御されリセットすれば位相が  $0^\circ$  に初期化される電圧制御正弦波発生器の例である。端子  $2_1$  の電圧によつて電流が制御される電流源 23, 24 が設けられ、これら電流源 23, 24 は電流スイッチ 25 により切替えられてコンデンサ 26 の一端及び電圧比較器 27, 28 に接続される。コンデンサ 26 よりの入力電圧が基準電圧  $V_r$  より高いと比較器 27 の出力が論理 1 になり、入力電圧が基準電圧  $-V_r$  より低いと比較器 28 の出力が論理 1 になり、その他は論理 0 を出力している。比較器 27 の出力により R フリップフロップ 29 がリセットされその出力は電流スイッチ 25 を制御する。もしフリップフロップ 29 の状態が 1 (セプト) のときスイッチ 25 は電流源 23 の側に接続され、コンデンサ 26 は充電される。フリップフロップ 29 の状態が 0 (リセプト) のときスイッチ 25 は電流源 24 の側に接続されコンデンサ 26 は放電される。端子

(10)

31に位相初期化入力パルスが加えられると位相初期化のためのスイツチ32が閉じてコンデンサ26が放電されるとともにR8フリップフロップ29がセットされる。このようにしてコンデンサ26の電位は三角波となりその周波数は電源23、24に対する端子21の制御電圧により変化する。また端子31より位相初期化入力にパルスを加えるとR8フリップフロップ29がセットされるとともにスイツチ32が閉じられてコンデンサ26が放電され、前記三角波の位相は0°に初期化される。コンデンサ26の三角波電圧は正低級変換器33により正低級へ変換されて端子34へ送出される。正低級変換器33としては通常折線近似が用いられる。以上に述べた電圧制御発振器の市販品の好例は米国インテール社製の電圧制御発振器IC(型番ICL8038)である。これには上記の位相初期化入力がないが、その他は第5図に示したものとほぼ同様の構成となっており、実験若干の付加回路によりこの発明の装置において利用できる。

(11)

ータ(balanced modulator)回路や各種のエキスパンダやコンプレッサ(例えばドルビー雑音低減方式やdbxノイズリダクションシステムにおいて使用されているものなど)が利用できる。

電圧制御パルス発生器12は周波数変換範囲は80~500 Hz程度と比較的狭いので各種の実現法が利用できる。簡単なものとしては電圧で制御できる電流源とコンデンサと放電回路を組合せたものがある。その一例を第7図に示す。端子11及び電源端子45との間の制御電圧がトランジスタ46のベースに与えられてそのコレクタ電流が制御され、そのコレクタ電流は抵抗器47を通じてコンデンサ48へ供給される。抵抗器47の両端はタイマ(米国シグネタイクス社製N555)49のトリガ端子、放電端子及びしきい値端子に接続され、このタイマ49よりパルスが端子51へ送出される。このほか各種のパルス変換回路も利用できる。

制御電圧の動きを平滑化フィルタは簡単なものとしては第8図に示すCR1次フィルタが利用で

(13)

乗算器あるいは可変利得増幅器あるいは可変減衰率減衰器を具体的に実現するには種々の方法がある。市販されている各種の乗算器は係数乗算器として利用できる。また第6図に示すような簡易な回路によつても利得を電圧または電流によつて制御できる増幅器が実現できる。図中Aは入力端子35を抵抗器36を通じてD&E T37のドレインに接続しD&E T37のソースを接地しグートに端子4から制御電圧を与え、D&E T37のドレイン電圧を出力端子38へ送出する。第6図Bは入力端子35とD&E T37のドレインを演算増幅器39の非反転入力端及び反転入力端にそれぞれ接続し演算増幅器39の出力端子と反転入力端との間に抵抗器41が接続される。第6図Cは第6図AのD&E T37の代りに発光ダイオード42及び光導電素子43のフォトカプラ44が用いられる。第6図Dは第6図Cの抵抗器36とフォトカプラ44とを入れ替えたものである。第6図Eは第6図BのD&E T37の代りにフォトカプラ44を用いた場合である。その他にもバランストモジュレ

(12)

きる。包絡線の発生は指数関数の場合は非常に簡単であり、第9図に示すようにパルス発生器12の出力パルスによりスイツチ52が瞬時のみ閉じられ、その瞬間端子16の電圧にコンデンサ53が充電されスイツチ52が開かれるとコンデンサ53の電荷は抵抗器54を通じて放電される。コンデンサ53の電圧は端子55に包絡線出力56として出力される。

以上説明したようにこの発明は従来の音声合成装置に比較して極めて簡単な構成を持ちハードウェアも容易に実現できる。しかも合成音声の品質は比較的良好であるという特徴を持つ。また音声合成フィルタをまったく使用せず、直接的に波형을合成するものであるため増幅ループの安定性や発振防止の問題を本質的に持たない。このためかなり雑な取扱いにも耐え得る。従つてこの発明の音声合成装置は安価な音声合成システムとして有用であり、その応用分野としては歌や空想などのアナウンスマシン、銀行のキャッシュディスプレイ、電話の号案内、外国語教育機器、音声出力つき

(14)

のほん成機、音声出力つきのコンピュータ機、その他各種の音声出力つきの通信機や民生機など多岐にわたることができる。なお上述はこの発明をアナログ信号処理により構成したが、デジタル処理により同様に構成することができる。この場合は各同一の機能部分は各正位設の系統について増分割多直利用することができる。

#### 4 図面の簡単な説明

第1図はこの発明による音声合成装置の実施例( $n=5$ としている)を示すブロック図、第2図はこの発明の他の実施例を示すブロック図、第3図は出力の包絡線をビブナ周期で変化させる例を示すブロック図、第4図は制御電圧を滑らかに変化させる例を示す図、第5図は位相初期化可能な電圧制御正弦波発振器の例を示す回路図、第6図は電圧制御可変利得増幅器及び可変減衰率減衰器の各種例を示す回路図、第7図は電圧制御可変周波数発生器の例を示す回路図、第8図は制御電圧の平滑器の例を示す回路図、第9図は包絡線発生器の例を示す回路図である。

1<sub>1</sub> ~ 1<sub>n</sub> : 電圧制御正弦波発振器、2<sub>1</sub> ~ 2<sub>n</sub> : 周波数制御電圧入力端子、3<sub>1</sub> ~ 3<sub>n</sub>、13 : 乗算器あるいは可変利得増幅器あるいは可変減衰率減衰器、4<sub>1</sub> ~ 4<sub>n</sub> : 振幅制御電圧入力端子、12 : 電圧制御パルス発生器、13 : 加算器、17 : 包絡線発生器、21、22 : 平滑化フィルタ。

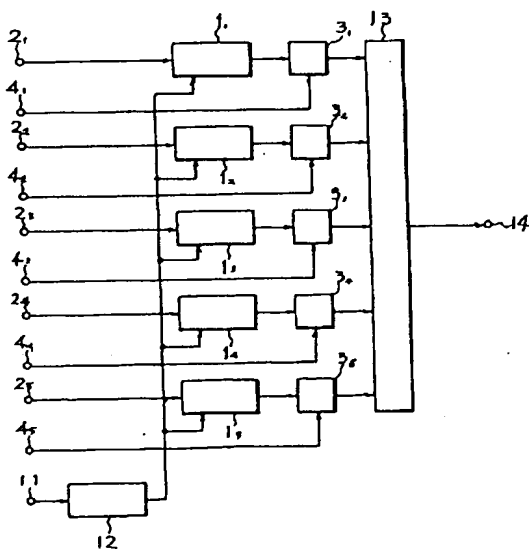
特許出願人 日本電信電話公社

代理人 岸 野 卓

(15)

(16)

第 1 図



第 2 図

